**Queried from NCBI Influenza Virus Resource**

* Human hosts,
* 16721 total sequences
* 47958 total (uncollapsed)

**Cleaned sequences, 16267 sequences remained**

**CD-Hit**-1570 remained